*Article*

# Estimating Subsurface Thermohaline Structure in the Tropical Western Pacific Using DO-ResNet Model

Xianmei Zhou [ID], Shanliang Zhu [ID], Wentao Jia and Hengkai Yao *

School of Mathematics and Physics, Qingdao University of Science and Technology, Qingdao 266061, China; zhouxianmei@mails.qust.edu.cn (X.Z.); zhushanliang@qust.edu.cn (S.Z.); jiawentao@mails.qust.edu.cn (W.J.)
* Correspondence: yaohengkai@qust.edu.cn

**Abstract:** Estimating the ocean's subsurface thermohaline information from satellite measurements is essential for understanding ocean dynamics and the El Niño phenomenon. This paper proposes an improved double-output residual neural network (DO-ResNet) model to concurrently estimate the subsurface temperature (ST) and subsurface salinity (SS) in the tropical Western Pacific using multi-source remote sensing data, including sea surface temperature (SST), sea surface salinity (SSS), sea surface height anomaly (SSHA), sea surface wind (SSW), and geographical information (including longitude and latitude). In the model experiment, Argo data were used to train and validate the model, and the root mean square error (RMSE), normalized root mean square error (NRMSE), and coefficient of determination ($R^2$) were employed to evaluate the model's performance. The results showed that the sea surface parameters selected in this study have a positive effect on the estimation process, and the average RMSE and $R^2$ values for estimating ST (SS) by the proposed model are 0.34 °C (0.05 psu) and 0.91 (0.95), respectively. Under the data conditions considered in this study, DO-ResNet demonstrates superior performance relative to the extreme gradient boosting model, random forest model, and artificial neural network model. Additionally, this study evaluates the model's accuracy by comparing its estimations of ST and SS across different depths with Argo data, demonstrating the model's ability to effectively capture the most spatial features, and by comparing NRMSE across different depths and seasons, the model demonstrates strong adaptability to seasonal variations. In conclusion, this research introduces a novel artificial intelligence technique for estimating ST and SS in the tropical Western Pacific Ocean.

**Keywords:** oceanography; machine learning; ocean thermohaline structure; remote sensing data; tropical Western Pacific

## 1. Introduction

Temperature and salinity are two key parameters of seawater, playing significant roles in understanding marine ecosystems, ocean dynamics, and climate change. Changes in temperature can impact various phenomena. These include marine heatwaves (periods of abnormally high sea surface temperatures), thermocline formation, and the evolution of El Niño [1–3]. Similarly, salinity variations are crucial for understanding the characteristics of the hydrological cycle, ocean circulation, and water mass formation [4–6]. Therefore, accurately estimating the subsurface temperature (ST) and subsurface salinity (SS) is vital for comprehending ocean dynamics, climate change, and predicting future climate scenarios.

However, accurately estimating ST and SS has been a significant challenge. Traditional methods, including forecast numerical simulations, data assimilation, dynamical models, and statistical models, have been widely used to estimate these parameters [7–12]. For instance, Wan et al. [13] analyzed the performance of two data assimilation methods: the ensemble Kalman filter and ensemble optimal interpolation. They used the hybrid coordinate ocean model for accurately predicting temperature and salinity in the Pacific system. The accuracy of the dynamic ocean model largely relies on observational data as

input. However, the spatiotemporal data provided by in situ observations are often sparse and discontinuous due to limitations inherent in monitoring systems such as buoys and Argo floats. These systems have limited coverage and are often constrained by deployment locations and operational durations. Data gaps can negatively impact data assimilation and forecasting accuracy. This can potentially lead to the omission of complex ocean dynamics [14], which may ultimately affect ST and SS estimation. Consequently, due to these limitations, classical methods struggle with challenges such as limited spatial coverage, low spatiotemporal resolution, and reduced accuracy.

In response to these challenges, remote sensing technology has rapidly developed over the past few decades. This technology offers an alternative approach to acquiring continuous and extensive sea surface data. Parameters such as sea surface temperature (SST), sea surface salinity (SSS), and sea surface height (SSH) have been accumulated. Many studies have demonstrated that subsurface properties can be characterized by related surface parameters [15–23]. For example, Chu et al. [16] demonstrated that the ocean interior temperature structure is closely related to SST. They showed that SST can be used to estimate the ST. Consequently, many researchers have combined ocean surface data with methods such as linear regression [17], least-squares regression [18,19], and empirical orthogonal functions (EOF) [20,23] to retrieve the three-dimensional structure of the ocean interior. For instance, Guinehut et al. [22] utilized satellite altimeter data and sea surface observations to derive global temperature and salinity fields through linear regression methods. Similarly, Maes et al. [21] reconstructed the subsurface salinity structure of the tropical Western Pacific by employing EOF with SSH and SST data. Although remote sensing technology provides continuous and extensive sampling, traditional methods such as data assimilation and numerical simulations of subsurface ocean variables are still challenging. These methods are often computationally intensive and do not always ensure high accuracy in estimations. Additionally, inherent differences between various data sources and observation methods can also impact the model's effectiveness. Given the limitations of traditional methods, researchers have increasingly adopted more advanced computational techniques to overcome these challenges.

In recent years, machine learning (ML) models have attracted significant attention in the field of oceanography. These models offer a promising alternative to traditional methods. ML models effectively explore latent relationships within the data. They estimate the internal physical structure of the ocean based on various observational datasets [24–27]. Early studies have employed machine learning models to retrieve the ocean's internal structure from multi-source surface data. For example, Ali et al. [28] utilized artificial neural networks (ANN) to reconstruct the vertical temperature profile of the Arabian Sea based on surface ocean data. This data included SST, SSH, wind stress, net radiation, and net heat flux. In another study, Wu et al. [29] integrated the self-organizing map (SOM) neural network algorithm with various satellite remote sensing datasets. They used this integration to reconstruct the subsurface temperature profile of the North Atlantic, demonstrating their model's superiority over traditional methods. Given the spatial correlation of data, Chen et al. [30] combined EOF analysis with the SOM method to estimate the subsurface temperature structure of the Northwest Pacific using sea surface data. EOF analysis simplifies and decomposes complex spatial data structures, while the SOM method captures nonlinear relationships within the data. The integration of these two methods is better suited for describing the nonlinear dynamic processes in the ocean. Additionally, Dong et al. [31] proposed a light gradient boosting machine-deep forest method to estimate the SS of the South China Sea. This approach showcased the potential of this approach in SS estimation. Su et al. [32–34] used various satellite source data to estimate global ocean subsurface temperature anomalies with machine learning methods such as random forest (RF), support vector regression, and other artificial intelligence methods. They also applied the extreme gradient boosting (XGBoost) model to estimate anomalies in both subsurface temperature and salinity across the global ocean [35]. These studies underscore the increas-

ing importance of ML models in advancing oceanographic research and enhancing the accuracy of ocean parameter estimation.

Building upon this foundation, more sophisticated deep learning (DL) models have been used to estimate ST and SS. These models further improve the estimation of oceanic parameters. Meng et al. [36] designed a deep neural network model based on convolutional neural networks (CNNs). This model uses multi-source satellite sea surface data to estimate the STA and subsurface salinity anomalies in the Pacific at a higher spatial resolution. By doing so, they significantly improve the resolution and accuracy of satellite observations in estimating internal oceanic parameters. Furthermore, Su et al. [37] combined sea surface remote sensing observations and Argo data to propose a bi-directional long short-term memory neural network method for predicting global ocean STA and subsurface salinity anomalies. This approach effectively learns significant temporal features of ocean variability, thereby enhancing the generalization capability of the prediction model. Cheng et al. [38] utilized the backpropagation neural network method to determine the ST in the North Pacific using SSH, SST, SSS, sea surface wind (SSW), and sea surface velocity obtained from remote sensing satellites. Mao et al. [39] proposed a dual-path convolutional neural network (DP-CNN) to reconstruct the ST and SS structures in the South China Sea using sea surface information. Their research demonstrated that DP-CNNs effectively reduce the issue of detail loss in CNN models. These DL-based research methods provide new avenues for the estimation of ST and SS.

Existing ML and DL models have shown the capability to estimate ST and SS in many regions. However, the tropical Western Pacific region is less documented. Additionally, their computational accuracy still requires improvement, especially in regions with complex dynamic processes. For instance, Wang et al. [40] proposed an enhanced ANN model based on multi-source sea surface data to estimate ST in the Western Pacific. However, the model's average root mean square error (RMSE) value reached as high as 0.55 °C. This lack of precision may stem from the model's limited capacity to extract and learn the spatiotemporal features of complex nonlinear systems. Therefore, there is considerable potential for improvement in both the model itself and its accuracy.

Motivated by the aforementioned discussions, this study aims to develop a novel method based on multi-source remote sensing data. The method will estimate the ST and SS simultaneously in regions with complex dynamic processes in the Western Pacific region. Research outcomes in this area are relatively scarce, with existing studies typically focusing on either ST or SS [40,41]. Our proposed method addresses both parameters simultaneously in a complex, dynamic environment. Consequently, this study proposes a new estimation method using the double-output residual network (DO-ResNet) model. The model aims to utilize multiple satellite remote sensing datasets, including SST, SSS, sea surface height anomalies (SSHA), SSW, and geographic information. These data will estimate the thermohaline structure in the tropical Western Pacific. Additionally, a series of estimation models are introduced for comparison in order to evaluate the proposed model. The remainder of this paper is organized as follows: Section 2 describes the study area and data. This section details the importance of the area studied and the multi-source remote sensing datasets used in the research. Section 3 outlines the research methods. It describes the construction and training process of the DO-ResNet model. Section 4 presents the research results. This section exhibits the ST and SS estimated by the DO-ResNet model and assesses the model's performance and accuracy through error statistics and comparative analysis. Section 5 summarizes and discusses the research findings. It outlines the main discoveries and contributions and explores the limitations of the methods as well as directions for future research.

## 2. Study Area and Data

### 2.1. Study Area

The strong El Niño event occurred in the Central and Eastern Pacific, significantly impacting the global climate [42,43]. As research has progressed, the scope of El Niño

studies has gradually expanded to the tropical Western Pacific region (Figure 1). As the region with the highest sea surface temperatures globally, the tropical Western Pacific contains the largest and warmest body of water in the world [44]. In particular, during the El Niño event, the Niño 4 region within this area normally maintains sea surface temperatures at or above the threshold throughout the year, reflecting its significant role in climate dynamics. Many scholars have found that variations in temperature and salinity in the tropical Western Pacific critically impact regional oceanic phenomena such as tropical cyclones, El Niño–Southern Oscillation (ENSO) events, the Pacific Decadal Oscillation, and monsoon variations [45–48]. To monitor these climate events, extensive observation systems, such as the Tropical Atmosphere and Ocean/Triangle Trans-Ocean Buoy Network (TAO/TRITON) array, are deployed in the region. This array provides valuable subsurface data [49]. However, due to a lack of comprehensive observations, our understanding of the spatiotemporal variations in temperature and salinity in the tropical Western Pacific remains limited, significantly constraining research on the thermohaline structure of this region. Therefore, accurately estimating ST and SS in the tropical Western Pacific is crucial for understanding the variability of ocean–atmosphere heat fluxes and analyzing the dynamic mechanisms of oceanic phenomena in this area.
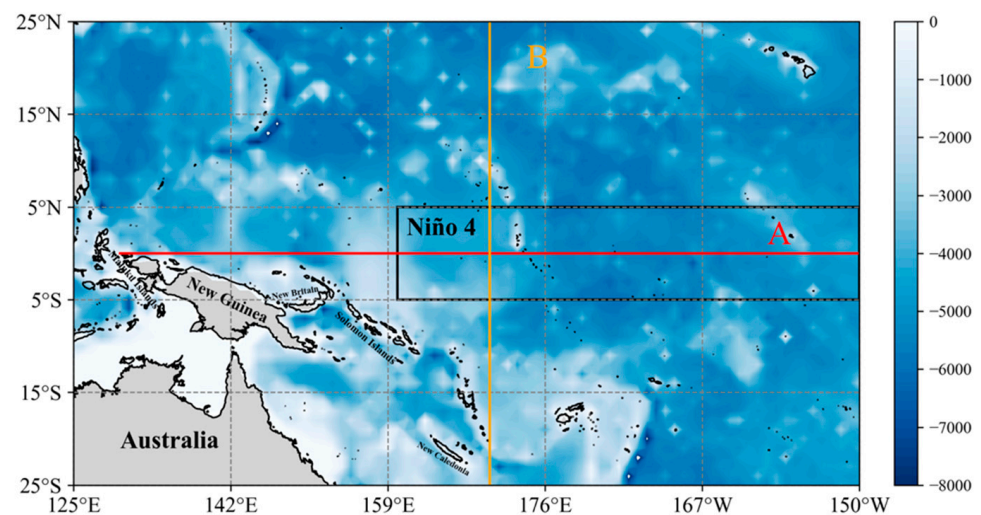


**Figure 1.** Bathymetry (unit: m) in the tropical Western Pacific, as well as locations of zonal section A (red line) and meridional section B (orange line) used in this study. The black box highlights the Niño 4 region.

*2.2. Data Source and Preprocessing*

The dataset for this study includes a series of sea surface remote sensing data and Argo gridded data for the tropical Western Pacific region, spanning from January 2010 to December 2020, as summarized in Table 1. The monthly SSS data utilized in this study are derived from the Soil Moisture and Ocean Salinity (SMOS) Level-3 product [50], featuring a spatial resolution of $0.25° \times 0.25°$. The monthly SST data are sourced from the National Oceanic and Atmospheric Administration (NOAA) [51], comprising interpolated observations from satellite radiometers, with a spatial resolution of $1° \times 1°$. The monthly SSHA data are obtained from the Archiving, Validation, and Interpretation of Satellite Oceanographic data (AVISO) project [52], with a spatial resolution of $0.25° \times 0.25°$. The SSW data, encompassing both the northward component (USSW) and the eastward component (VSSW), are acquired from the Cross-Calibrated Multi-Platform (CCMP) product [53], with a spatial resolution of $0.25° \times 0.25°$. To evaluate the model's performance, this study utilizes the Argo gridded data product (also referred to as Argo observations hereafter) developed by Roemmich and Gilson as the label data [54]. This product has undergone multi-level quality controls and interpolation procedures applied to the raw observational Argo floats collected from the global oceans. All Argo drifting floats designed to collect

ocean subsurface parameters are organized under the international Argo program, which stands for Array for Real-time Geostrophic Oceanography [55]. The resulting dataset provides global gridded ocean temperature and salinity with a spatial resolution of $1° \times 1°$, encompassing 58 standard depth levels ranging from 5 m to 1975 m. All datasets utilized in this study are detailed in Table 1.

**Table 1.** Summary of data used in this study.

| Index | Contents | | | | |
|---|---|---|---|---|---|
| Study Area | Tropical Western Pacific (25° S–25° N, 125° E–150° W) | | | | |
| Data | SSS | 2010–2020 | SMOS | | Input |
| | SST | 2010–2020 | NOAA | | |
| | SSHA | 2010–2020 | AVISO | | |
| | SSW | 2010–2020 | CCMP | | |
| | ST | 2010–2020 | Argo | | Label |
| | SS | 2010–2020 | Argo | | |
| Resolution | monthly | | | $0.5° \times 0.5°$ | |

To ensure the consistency and accuracy of the model, this study employs data spanning from January 2010 to December 2020. All datasets were processed into monthly averages and linearly interpolated to a resolution of $0.5° \times 0.5°$ to guarantee uniform temporal and spatial coverage. It is important to note that the use of monthly temporal resolution limits the ability to detect rapid and subtle changes in oceanic and atmospheric conditions that precede the El Niño event [56], and the interpolation process may affect the accuracy of the model. To address this, we performed a detailed variance analysis to quantify the impact of interpolation on the original data. The results showed minimal impact on the data's integrity. Data points with missing parameters within the study region were excluded from the analysis. For the model input, each data point within the region was considered a central point, with the 81 data points within a $9 \times 9$ grid surrounding the central point serving as a two-dimensional image representing that central point. Furthermore, to accelerate model convergence during training, all remote sensing and Argo data were normalized using their respective means and standard deviations.

## 3. Methods

### 3.1. The DO-ResNet Model

Residual neural network (ResNet) [57] is a deep convolutional neural network architecture for handling complex image recognition and classification tasks in the nonlinear case, which consists of multiple convolutional layers and residual blocks. Each residual block contains several convolutional layers and uses shortcut connections to add the original input to the output of the convolutional layers, thus preserving the information from the input data. Compared to traditional CNNs, ResNet introduces shortcut connections while retaining the characteristics of CNNs, effectively addressing the vanishing gradient problem in deep neural networks and demonstrating exceptional intelligent characteristics. In recent years, people have applied ResNet in various fields to solve practical problems, such as drug-induced liver injury prediction [58], coral reef semantic segmentation [59], sea ice detection [60], and others [61,62].

In view of the close relationship between the sea surface parameters and the subsurface thermohaline structure of the ocean, a double-output ResNet (DO-ResNet) model is proposed in this study, which can simultaneously estimate ST and SS based on various sea surface parameters observed by satellites. The double-output design allows us to extract features of both temperature and salinity through a single model, resulting in two outputs and improving the model's efficiency. The structure of the improved DO-ResNet model is

illustrated in Figure 2a. Except for the dual output and the first convolutional layer, the rest of the structure is the same as the original ResNet model. Firstly, the input data passes through a convolutional layer for preliminary feature extraction. This convolutional layer uses a $3 \times 3$ kernel, which effectively captures local features while reducing the impact of islands on the ocean temperature and salinity information at the ocean–island boundaries. Secondly, after deep feature extraction through the residual blocks, the data is flattened and fed into a fully connected layer for integration. Finally, the model estimates temperature and salinity through two independent output layers.
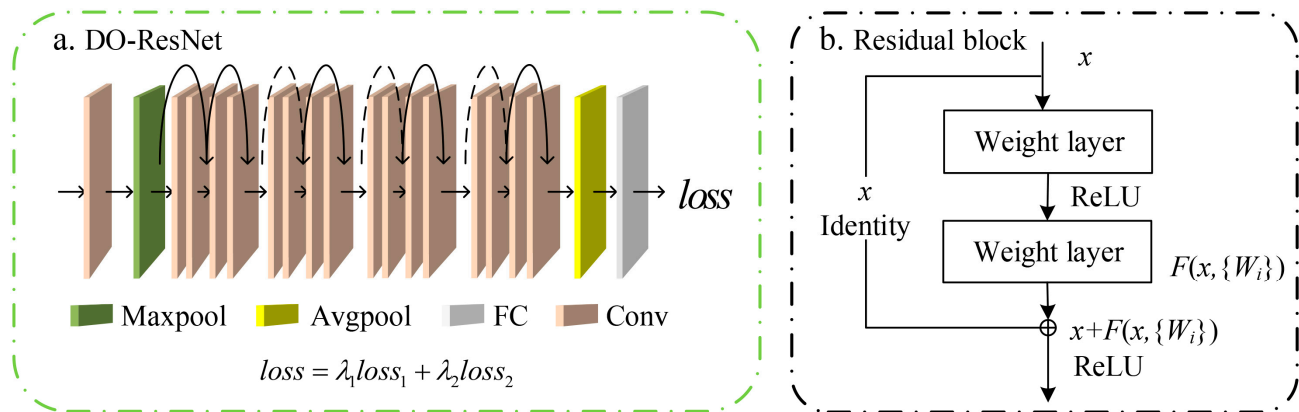


**Figure 2.** The structure of DO-ResNet model, with the solid and dashed lines representing two different types of shortcut connections (**a**) and schematic diagram of a residual block (**b**), including two convolutional layers, two activation functions, and a shortcut connection.

The structure of each residual block is shown in Figure 2b, consisting of two $3 \times 3$ convolutional layers, two rectified linear unit (ReLU) [63] activation functions, and a shortcut connection. In each residual block, the input data $x$ is processed sequentially through the first convolutional layer, the ReLU activation function, and the second convolutional layer to obtain the output $F(x, \{W_i\})$. To preserve the original information of the input data, the input data $x$ is directly connected to the output of the second convolutional layer via the shortcut connection. Finally, after passing through the ReLU activation function, the final output $y$ of the residual block is generated. It is important to note that when the input and output dimensions are the same, the shortcut connection is represented by a solid line (as shown in Figure 2a), and the propagation process is as described in Equation (1):

$$y = x + F(x, \{W_i\}) \tag{1}$$

Here, $x$ and $y$ are the input and output vectors of the layers considered. $F(x, \{W_i\}) = \sigma(W_1 x)$ represents the learned residual mapping, where $\sigma$ denotes the ReLU activation function. Similarly, shortcut connections that span feature maps of different dimensions are represented by dashed lines (as shown in Figure 2a). In this case, a $1 \times 1$ convolutional layer is used to perform dimensionality upscaling on the input data $x$, and the propagation process is as described in Equation (2):

$$y = W_s x + F(x, \{W_i\}) \tag{2}$$

Here, $x$, $y$, and $F(x, \{W_i\})$ have the same meanings as in Equation (1), and $W_s$ represents the identity mapping obtained by processing the input $x$ on the shortcut connection branch. This type of connection effectively reduces the loss of input information during propagation, preserves the integrity of the input data, and simultaneously lowers the risk of gradient vanishing, thereby enhancing the stability and performance of the neural network.

### 3.2. Experimental Setup

This study utilizes multi-source satellite observational data to develop a DO-ResNet model for estimating ST and SS within the tropical Western Pacific region. The research workflow is illustrated in Figure 3 and is divided into three main stages. The first stage involves the collection and preprocessing of raw data. We collected five sea surface parameters: SST, SSS, SSHA, USSW, and VSSW, as well as longitude (LON) and latitude (LAT) from multiple databases. These datasets were then preprocessed to create a training and testing dataset comprising gridded data points over 132 months. ST and SS data derived from Argo data served as the training and testing labels. The second stage is model training. The monthly average data from January 2010 to December 2019 were used as training data to train the DO-ResNet model. During this phase, we selected the mean squared error (MSE) as the loss function. MSE is a smooth and continuous function that facilitates efficient gradient calculation and optimization. Additionally, its sensitivity to outliers helps the model reduce errors effectively [64,65]. To better estimate both temperature and salinity simultaneously in our model, a composite loss function incorporating the MSE of these two variables is designed, as shown in Equation (3):

$$loss = \lambda_1 loss_1 + \lambda_2 loss_2 \tag{3}$$
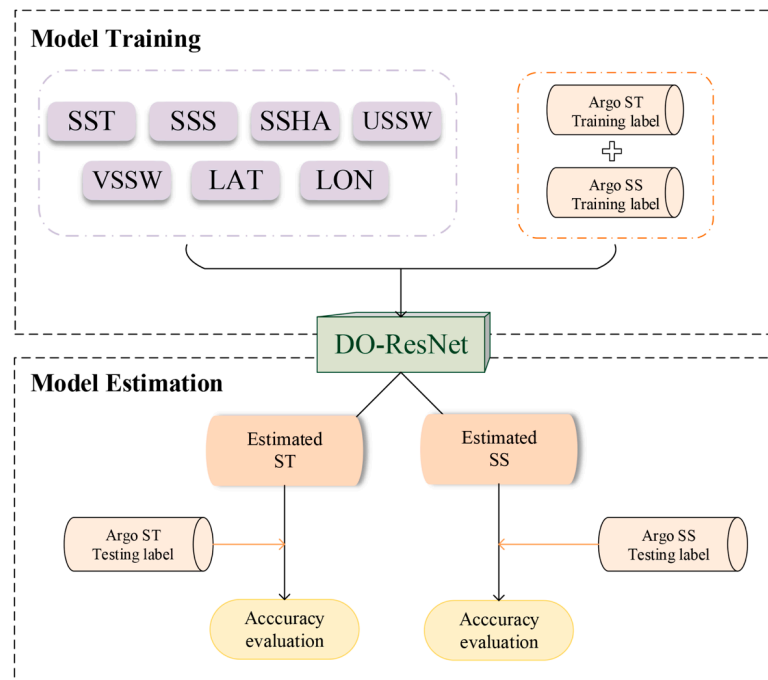


**Figure 3.** Flowchart of the DO-ResNet model for estimating ocean subsurface thermohaline structure in the tropical Western Pacific.

Here, the coefficients $\lambda_1$ and $\lambda_2$ are weights that balance the contributions of the temperature and salinity losses. The $loss_1$ and $loss_2$ represent the MSE for the temperature and salinity estimates, respectively, with their calculation formulas given by Equations (4) and (5).

$$loss_1 = \frac{1}{N}\sum_{i=1}^{N}\left(T_i - \hat{T}_i\right)^2 \tag{4}$$

$$loss_2 = \frac{1}{N}\sum_{i=1}^{N}\left(S_i - \hat{S}_i\right)^2 \tag{5}$$

Here, $T_i$ and $S_i$ are the actual temperature and salinity values at the *i*-th data point, respectively, while $\hat{T}_i$ and $\hat{S}_i$ are the model-estimated temperature and salinity values at that data point. $N$ is sample number.

By adjusting the weight coefficients $\lambda_1$ and $\lambda_2$, this composite loss function measures the training loss for both temperature and salinity, thereby optimizing the model's overall performance. To reduce computational cost and based on the quasi-linear relationship between ocean temperature and salinity in the T−S diagram, we set both $\lambda_1$ and $\lambda_2$ to 0.5 [66]. However, it is important to note that this simplification might prevent the model from achieving optimal performance, as there could be a more effective set of coefficients. Additionally, to prevent overfitting, we employed random sampling of data points and an early stopping strategy. The early stopping strategy halts training when the model's accuracy does not improve over several consecutive iterations. Moreover, we used a grid search method to identify the optimal parameter combination for the DO-ResNet model, as summarized in Table 2.

**Table 2.** Parameter values of DO-ResNet models.

| Estimation Models | Parameter Values |
|---|---|
| DO-ResNet | convolutional layer : size $= 3 \times 3$, stride $= 1$; adaptiveavgpool2d : output_size $= 1 \times 1$; loss function: mse; optimizer: radam; learning rate: 0.02; reducelronplateau: mode = 'min', factor = 0.1, patience = 10; batch size: 2048; activation function: relu; batchnorm2d; validation frequency: per epoch earlystopping: patience = 7, verbose = False, delta = 0 |

Finally, the model testing is the third stage, as shown in Figure 3. The test data from each month of 2020 were input into the model to obtain the prediction results. The performance of the DO-ResNet model was evaluated using the RMSE and determination coefficient ($R^2$). RMSE, which measures the difference between estimated and observed values, The $R^2$ metric is employed to evaluate the model's fitting performance, thereby assessing its estimation capability. It should be noted that, as a deep learning model, the accuracy of DO-ResNet is highly sensitive to the quality of the original input data.

## 4. Results and Discussion

### 4.1. Identification of Input Variable

A key issue in estimating ST and SS in the tropical Western Pacific is the selection of model input variables. Previous studies have indicated that incorporating SSW and geographic information can enhance the accuracy of subsurface temperature and salinity estimations [37,67,68]. To further investigate the impact of SSW and geographic information on temperature and salinity estimations in this region, this study designed experiments with three different combinations of input parameters (Cases 1, 2, and 3) to compare the results estimated by DO-ResNet with different training inputs. Table 3 presents the parameter combinations for the three experimental cases. In Case 1, SSS, SST, and SSHA were selected as input parameters. In Case 2, both USSW and VSSW were added to these parameters. In Case 3, geographic information (LON and LAT) was added to the input parameters of Case 2.

**Table 3.** Presents the optimized combinations of model parameters.

| Experiments | Training Methods |
|---|---|
| Case 1 (3 parameters) | ST (SS) = Ensemble (SST, SSS, SSHA) |
| Case 2 (5 parameters) | ST (SS) = Ensemble (SST, SSS, SSHA, USSW, VSSW) |
| Case 3 (7 parameters) | ST (SS) = Ensemble (SST, SSS, SSHA, USSW, VSSW, LON, LAT) |

Figure 4 shows the vertical distribution of RMSE and $R^2$ for the three experimental cases. The results indicate that three experiments are able to estimate the vertical distribution of ST (SS), with the main error centered on the depth of 150–200 m, which may be an effect of the thermocline [69]. However, the inclusion of SSW and geographic information in Case 3 improves the estimation accuracy of the DO-ResNet model in the tropical Western Pacific Ocean compared to Cases 1 and 2, as evidenced by smaller RMSE values and larger $R^2$ values. Specifically, the vertically averaged RMSE and $R^2$ for estimating the ST (SS) were 0.51 °C (0.07 psu) and 0.84 (0.92) in Case 1, 0.46 °C (0.07 psu) and 0.87 (0.93) in Case 2, 0.35 °C (0.05 psu), and 0.91 (0.96) in Case 3, respectively. The DO-ResNet model in Case 3 exhibited significantly lower RMSE values across all depths compared to the results from Cases 2 and 1, and the $R^2$ values were also notably higher. These results show that the DO-ResNet model for Case 3 can better fit the nonlinear relationship between the ST (SS) in the tropical Western Pacific and the input parameters. Similarly, Qi et al. found that incorporating geographic information improved model accuracy in their study of Pacific ST [70], a finding that aligns with the improvements observed in Case 3 of our study. However, including latitude and longitude information may limit the model's generalizability. Furthermore, the model's accuracy needs further enhancement in the dynamically complex thermocline region and in areas with strong local gradients.
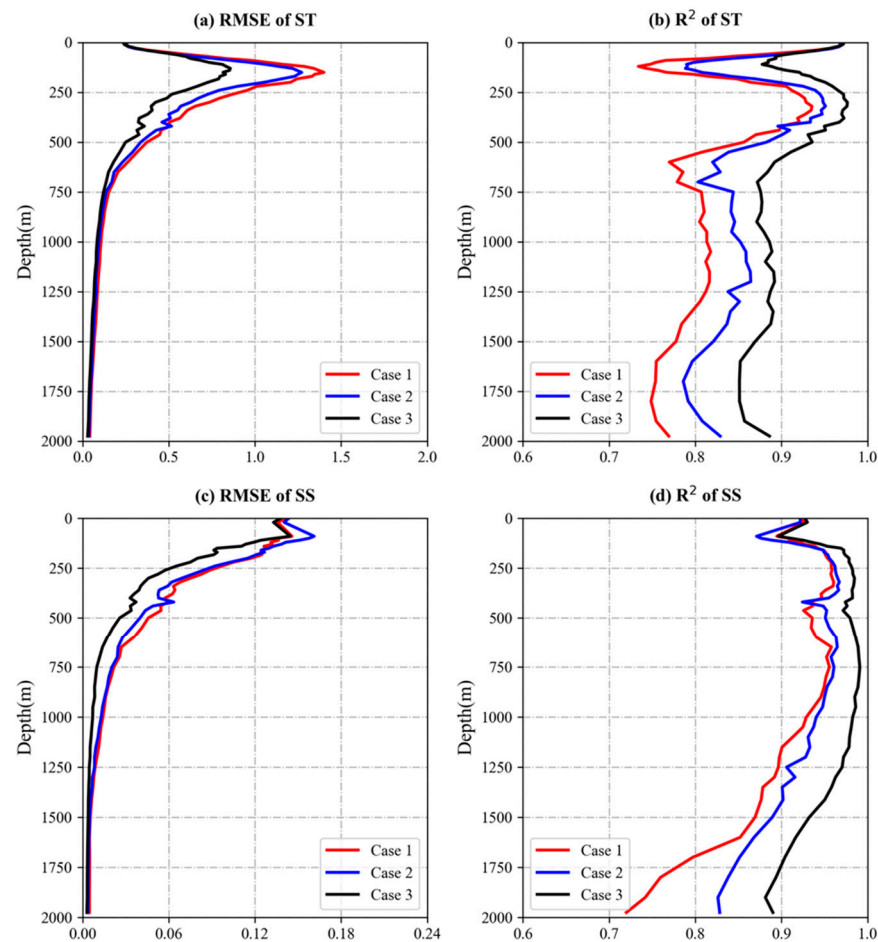


**Figure 4.** The estimation accuracy of the annual mean ST (**a**,**b**) and SS (**c**,**d**) at different depths by the DO-ResNet model based on RMSE and $R^2$ in different cases (Cases 1, 2, and 3 shown in red, blue, and black lines, respectively) in 2020.

*4.2. Accuracy Comparison between the DO-ResNet Model and Other Models*

After determining the optimal parameter combination, this study compared the performance of the DO-ResNet model with three other models: XGBoost (extreme gradient boosting, a machine learning technique known for its high performance and scalability, version 1.3.3) [35], RF (an ensemble learning method that constructs multiple decision trees for classification and regression) implemented with sklearn module (version 0.23.2) in Python [33], and ANN (a class of models inspired by the human brain for pattern recognition and prediction) [40] for estimating ST and SS in the tropical Western Pacific using 2020 data. The parameter values for the models are detailed in Table 4.

**Table 4.** Parameter values of different estimation models.

| Models | Parameter Values |
|---|---|
| XGBoost | eta = 0.02, min_child_weight = 2.0, max_depth = 5, subsample = 0.8 |
| RF | min_samples_split = 100, min_samples_leaf = 20, max_depth = 8, random_state = 10 |
| ANN | number of neural network layers = 3, learning rate = 0.002, number of neurons per layer = 30, loss function = MSE |

Figure 5a,b illustrates the vertical distribution of RMSE and $R^2$ for ST estimates by the four models. The findings reveal that the RMSE of ST estimates for all models increases from the sea surface, peaks around 200 m depth, and then decreases with further depth. In contrast, $R^2$ values show an inverse pattern, decreasing from the surface to 200 m, peaking around 300 m, and then fluctuating. The DO-ResNet model outperforms the other models with consistently lower RMSE and higher $R^2$ values across various depths. For SS estimates, shown in Figure 5c,d, a similar trend is observed: RMSE increases and then decreases, while $R^2$ decreases initially and peaks at 750 m, and then gradually decreases. At depths of 150–200 m, the DO-ResNet model's RMSE for SS is relatively higher, and its $R^2$ is relatively lower, possibly due to the thermocline, which complicates accurate estimation [69]. The model may struggle to capture the complex dynamics within this depth range. Despite this, the DO-ResNet model's average RMSE and $R^2$ for ST (SS) estimates across all depths are 0.35 °C (0.06 psu) and 0.91 (0.96), while the overall average RMSE and $R^2$ for XGBoost, RF, and ANN are 0.50 °C (0.07 psu) and 0.87 (0.93), 0.51 °C (0.06 psu) and 0.86 (0.93), and 0.43 °C (0.06 psu) and 0.90 (0.94), respectively. For comparison, the average RMSE value for ST estimated by Wang et al. using the neural network model was 0.55 °C, which is higher than that of the DO-ResNet model, indicating inferior performance [40]. These results clearly illustrate that the DO-ResNet model surpasses the other models in performance, as evidenced by its smaller RMSE and higher $R^2$ values. Therefore, the DO-ResNet model is highly effective in accurately estimating ST and SS in the tropical Western Pacific.
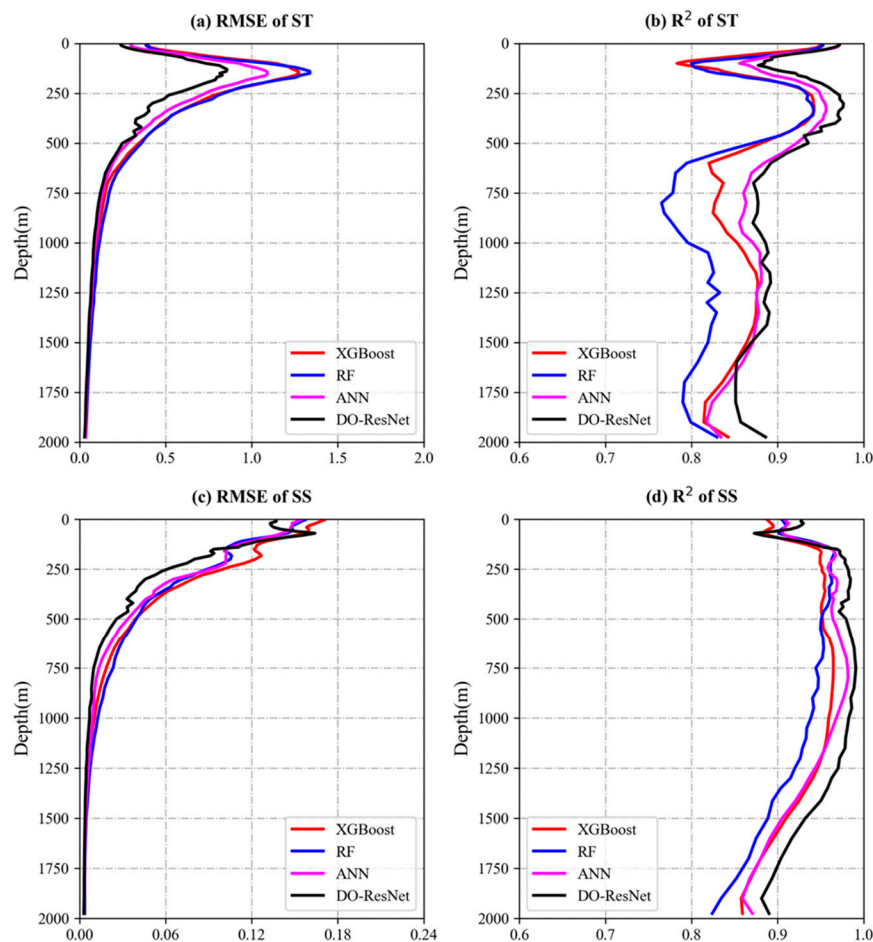
**Figure 5.** Comparison of the DO-ResNet model (black lines) with three machine learning models (XGBoost: red lines, RF: blue lines, ANN: magenta lines) for the annual mean ST and SS estimation at different depths based on RMSE (**a**,**c**) and $R^2$ (**b**,**d**) in 2020.

*4.3. Vertical Performance Evaluation of the DO-ResNet Model*

This section provides a comprehensive evaluation of the DO-ResNet model's performance in estimating ST and SS from various perspectives. For a comprehensive assessment of the vertical performance of the DO-ResNet model, we analyzed two representative sections using annual average data for 2020: a zonal section (section A) along the equator from 130° E to 150° W, and a meridional section (section B) along 170° E from 25° S to 25° N. The locations of these sections in the tropical Western Pacific are shown in Figure 1. The result is shown in Figures 6 and 7.

Figure 6a,b presents the spatial distribution of ST along the zonal section A as observed by Argo data and estimated by the DO-ResNet model, respectively. Similarly, Figure 6d,e displays the vertical distribution of ST along section B. The DO-ResNet model exhibits strong agreement with the Argo observations, accurately capturing the vertical distribution of ST. The 20 °C isotherm highlights a pronounced vertical gradient in both observed and estimated temperatures at the surface and within the thermocline, consistent with findings from previous studies regarding thermocline depth [71]. From the surface to the thermocline depth, seawater temperature decreases rapidly. Beyond the thermocline, the temperature decline is more gradual, reaching a stable state at depths below 500 m.
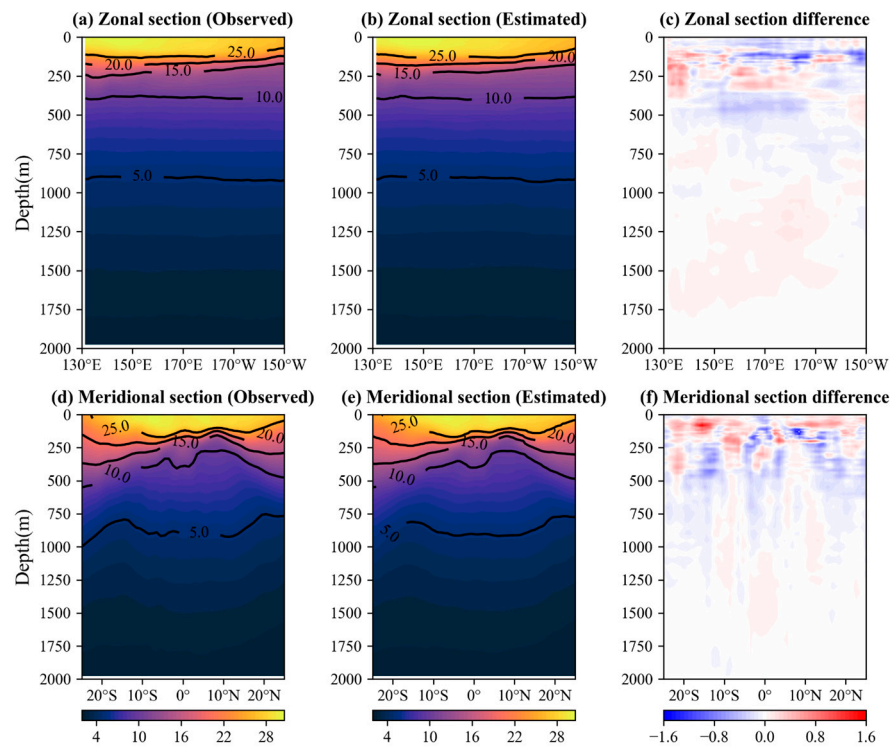
**Figure 6.** The 2020 annual mean Argo data (**a**,**d**), DO-ResNet model estimated (**b**,**e**), and their differences (Argo observations minus DO-ResNet estimates; (**c**,**f**)) of ST (unit: °C) along zonal and meridional sections.
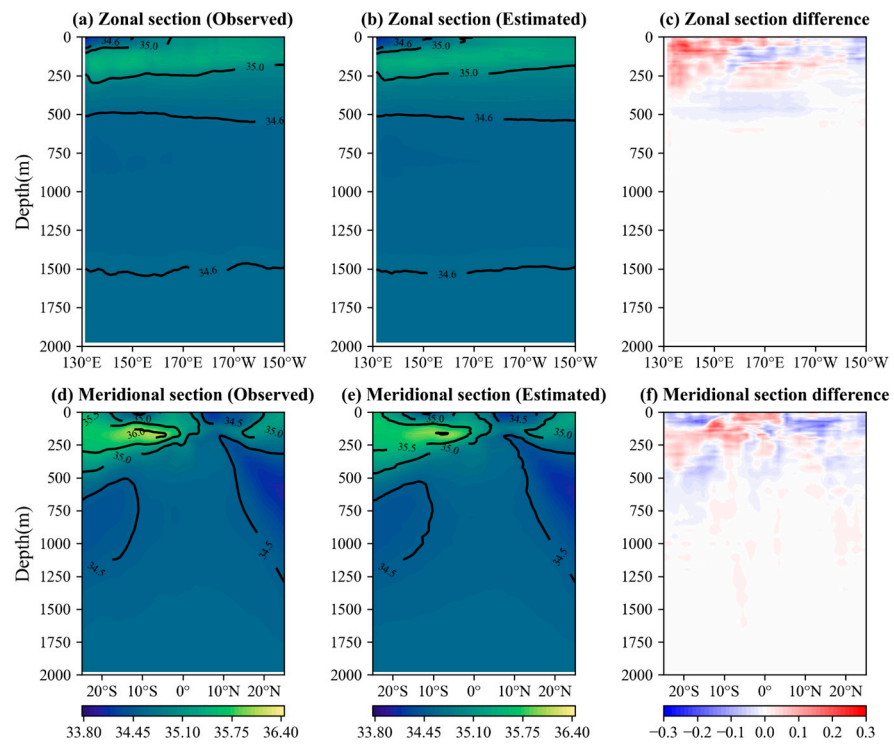


**Figure 7.** The 2020 annual mean Argo data (**a**,**d**), DO-ResNet model estimated (**b**,**e**), and their differences (Argo observations minus DO-ResNet estimates; (**c**,**f**)) for SS (unit: psu) along zonal and meridional sections.

Figure 6c,f displays the difference between the Argo observations and the DO-ResNet model estimates of ST along the section (Argo observations minus DO-ResNet estimates). The differences are generally minor, with more significant discrepancies primarily in the shallow layers. For instance, for section A, at depths shallower than 300 m between 130° E and 160° E, the model estimates are approximately 0.8 °C lower than the observations, while between 160° E and 150° W, there are regions where the model estimates are higher than the observations. Similarly, Figure 6f shows that larger differences (exceeding 0.5 °C) are primarily located in the shallow layers (50–500 m), likely due to the complex circulation system in the equatorial Pacific region [70].

Similarly, the DO-ResNet model's estimates of SS align well with Argo observations along both zonal and meridional sections, effectively reproducing the distribution characteristics of SS. Figure 7a,b illustrates the vertical distribution of SS along the zonal section, while Figure 7c depicts the differences between Argo observations and the DO-ResNet estimates. In the upper ocean layers, specifically between 50 and 100 m, the maximum salinity difference in certain regions (130° E to 150° E) reaches 0.23 psu, which may be influenced by the New Guinea Island and Maluku Islands. As depth increases, the DO-ResNet model's estimated SS stabilizes, with the difference between the estimated and observed salinity being less than 0.14 psu. Similarly, Figure 7d,e shows the vertical distribution of SS along the meridional section, while Figure 7f shows the differences between the two. With increasing depth, the discrepancies between the DO-ResNet model's estimates and the Argo observations along the meridional section gradually decrease. These results confirm the effectiveness of the DO-ResNet model in estimating SS in the tropical Western Pacific.

To evaluate the performance of the DO-ResNet model at different depths, Figures 8 and 9 present the annual mean ST and SS at six different depths (50 m, 200 m, 500 m, 1000 m, 1500 m, and 1900 m) estimated by the DO-ResNet model for 2020. The model's performance is evaluated by comparing the differences between the Argo data and the DO-ResNet estimates.

Figure 8 shows the spatial distribution of ST estimated by the DO-ResNet model at various depths, demonstrating a high degree of consistency with the observed ST. At 50 m depth, Argo observations indicate that the highest temperatures in the tropical Western Pacific occur at the equator, with sea surface temperatures gradually decreasing towards both poles, and a distinct thermal front appearing near 5° S. The DO-ResNet model accurately captures this significant thermal feature. At 200 m depth, the differences between the DO-ResNet model's estimated ST and the Argo observed ST range from −1.70 °C to 1.55 °C, which is greater than the differences at 50 m depth, likely due to the thermocline's influence. As depth increases, the temperature stabilizes, and below 1000 m, the differences between the temperatures from Argo data and DO-ResNet model estimates are minimal, ranging from −0.20 °C to 0.26 °C. These results confirm that the DO-ResNet model is highly accurate in estimating ST in the tropical Western Pacific.

Similarly, Figure 9 illustrates the spatial distribution of SS estimated by the DO-ResNet model in comparison with Argo observations. The results reveal that the SS values estimated by the DO-ResNet model show a high degree of consistency with the Argo observations, effectively reproducing the distribution characteristics of SS. In the upper ocean layers (50 m and 200 m), the salinity differences across most regions range from −0.29 psu to 0.26 psu. As depth increases, SS becomes more stable, and the differences between the salinity estimated by the DO-ResNet model and that provided by Argo are less than 0.1 psu. These findings confirm the reliability of the DO-ResNet model for estimating SS in the tropical Western Pacific.
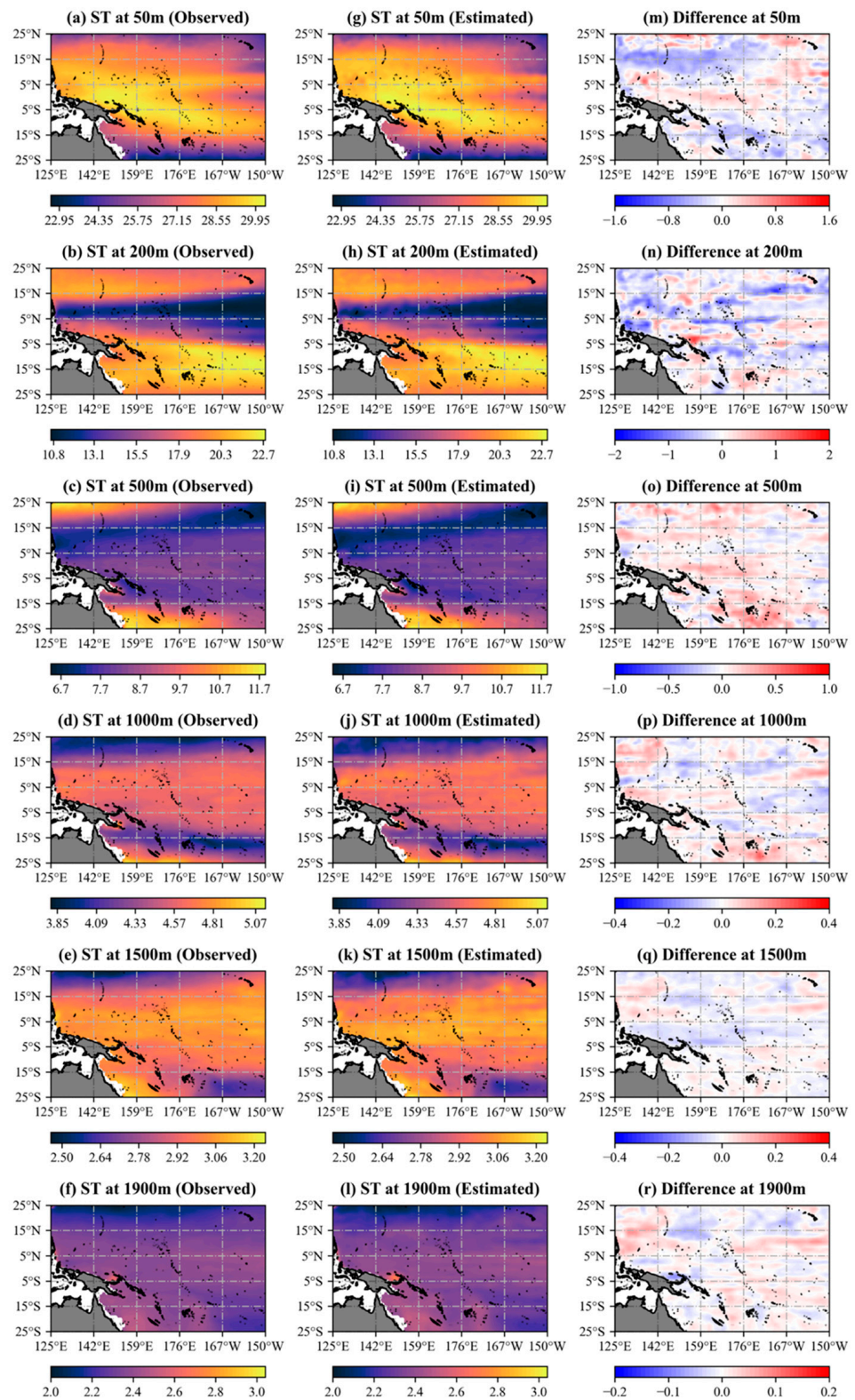
**Figure 8.** Annual averaged ST (unit: °C) from Argo observation (**a**–**f**) and DO-ResNet estimation (**g**–**l**) and their differences (Argo observations minus DO-ResNet estimates, (**m**–**r**)) at different depths (50 m, 200 m, 500 m, 1000 m, 1500 m, and 1900 m) in 2020.
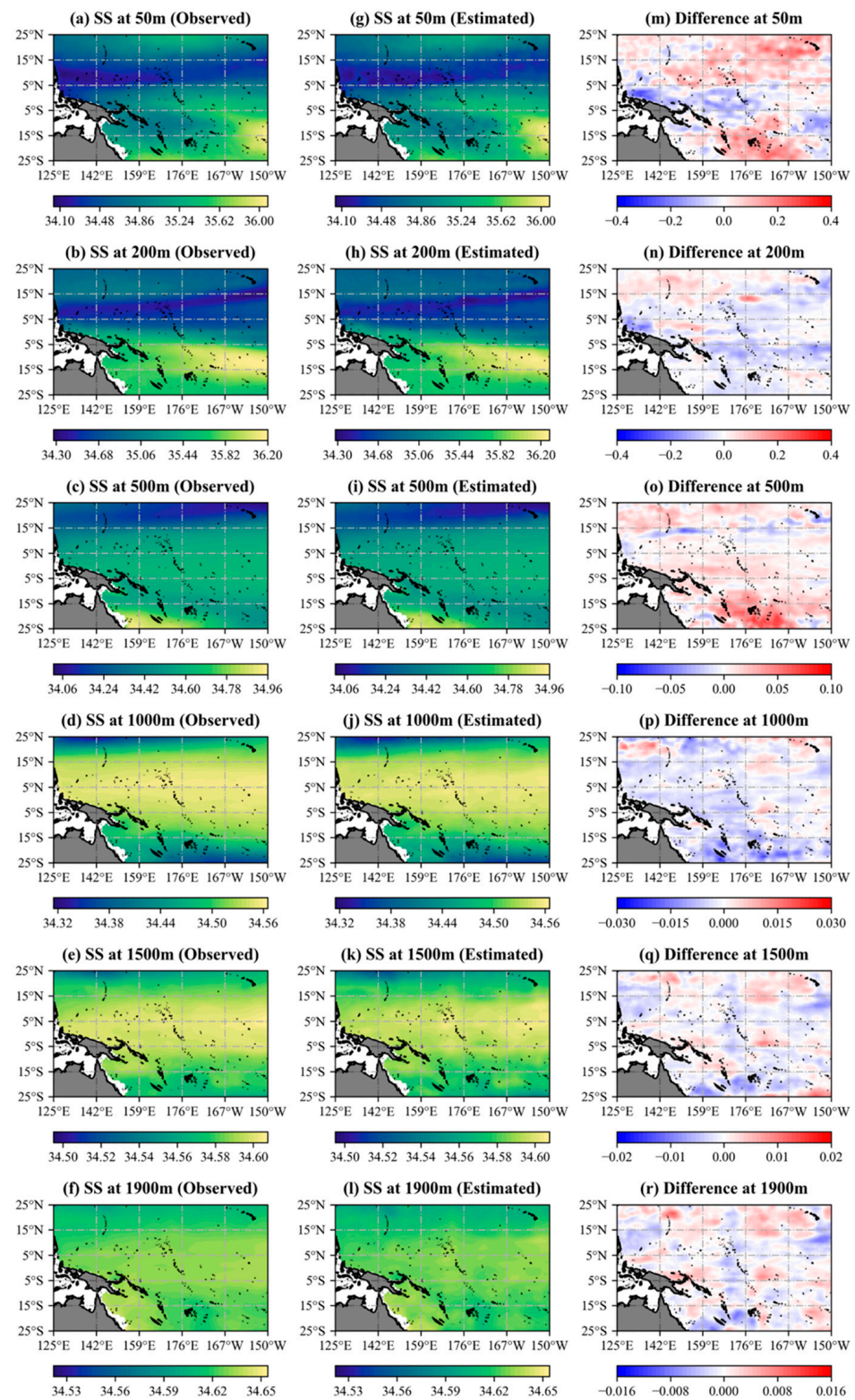
**Figure 9.** Annual averaged SS (unit: psu) from Argo observation (**a–f**) and DO-ResNet estimation (**g–l**) and their differences (Argo observations minus DO-ResNet estimates, (**m–r**)) at different depths (50 m, 200 m, 500 m, 1000 m, 1500 m, and 1900 m) in 2020.

The temperature–salinity (T−S) diagram is a valuable tool for studying the characteristics of water masses in the equatorial Western Pacific [72]. We selected February and August of 2020 to represent winter and summer, respectively. Using Argo data and model estimates for these months, we constructed T−S diagrams for the Western Pacific, as shown in Figure 10. The figure clearly demonstrates that both the Argo data and model results consistently capture the features of North Pacific Subtropical Underwater (NPSTUW), North Pacific Subtropical Mode Water (NPSTMW), North Pacific Central Water (NPCW), North Pacific Intermediate Water (NPIW), South Pacific Subtropical Mode Water (SPSTMW), South Pacific Central Water (SPCW), and Antarctic Intermediate Water (AAIW) during both summer and winter in the tropical Western Pacific [66].
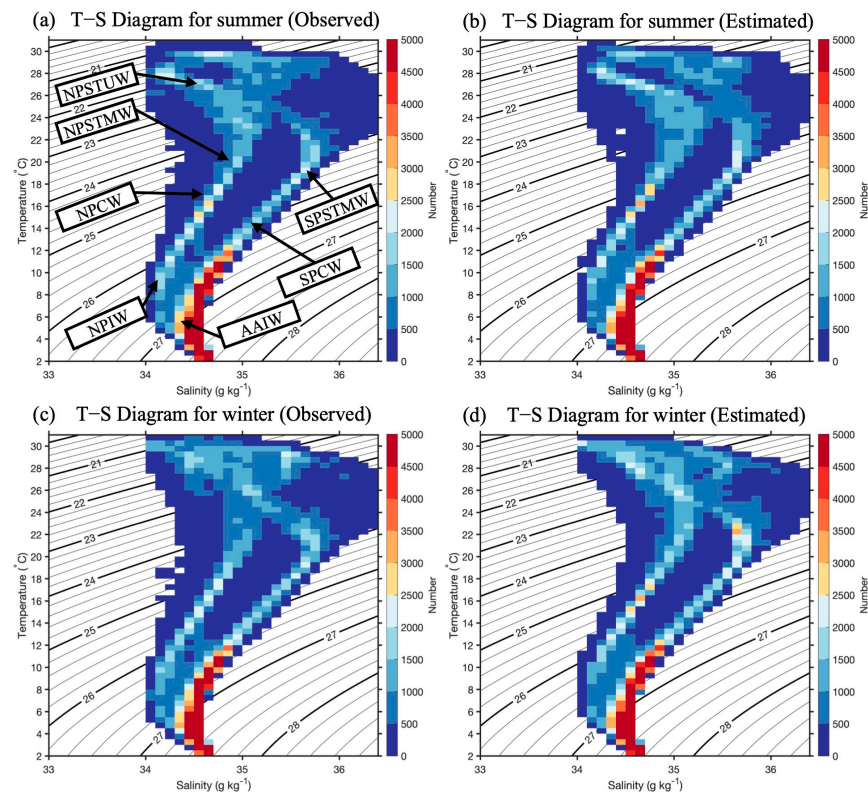


**Figure 10.** T−S diagrams of the Western Pacific region for winter (February 2020) and summer (August 2020), based on Argo data (**a**,**c**) and the DO-ResNet model results (**b**,**d**). Acronyms: NPCW, North Pacific Central Water; SPCW, South Pacific Central Water; NPSTUW, North Pacific Subtropical Underwater; NPSTMW, North Pacific Subtropical Mode Water; SPSTMW, South Pacific Subtropical Mode Water; NPIW, North Pacific Intermediate Water; AAIW, Antarctic Intermediate Water.

### 4.4. Seasonal Performance of the DO-ResNet Model

In this section, the months of February, May, August, and November of 2020 have been chosen to represent winter, spring, summer, and autumn, respectively, to investigate the model's performance across different seasons. The model's performance was evaluated at 22 different depths (ranging from 30 m to 1900 m) across these four seasons. To ensure comparability across depths, the RMSE values were divided by the standard deviation of the observed Argo temperature and salinity values at each depth, resulting in the NRMSE. The NRMSE and $R^2$ values for different seasons and depths are illustrated in Figure 11.
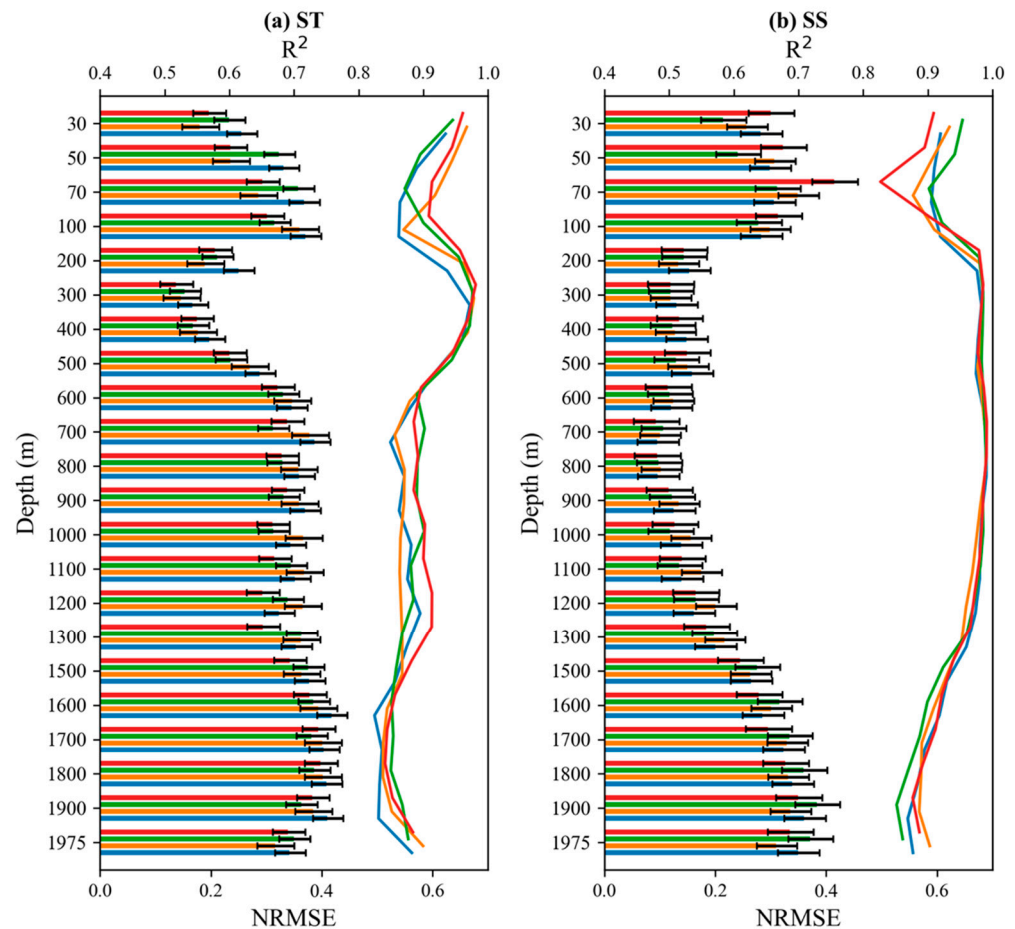
**Figure 11.** Seasonal performance of the DO-ResNet model for ST (**a**) and SS (**b**) estimation at different depths in the tropical Western Pacific in 2020. The data are presented with error bar representing 95% confidence interval. Blue indicates February (winter), orange indicates May (spring), green indicates August (summer), and red indicates November (autumn). The histograms display the NRMSE, while the lines display $R^2$.

Figure 11a presents the vertical distribution of NRMSE and $R^2$ values of the ST estimated by the DO-ResNet model across different seasons and depths. The NRMSE across different seasons exhibits a downward trend followed by an upward trend, with a turning point occurring at depths of approximately 200 to 400 m. Beyond 600 m, NRMSE rises with depth slowly, reflecting the model's weaker performance in deeper layers, making it harder to estimate deeper phenomena with surface data. Additionally, the higher NRMSE values at depths shallower than 100 m could be related to the complex dynamic processes occurring in the upper ocean. This pattern is consistent with previous studies on model-estimated ST in other regions [73]. The DO-ResNet model demonstrates robust performance in estimating ST across different seasons. The maximum NRMSE values for February and May are 0.40 and 0.42, respectively, while for August and November, they are 0.40 and 0.39, respectively, indicating better model performance in winter. The estimation accuracy of the DO-ResNet model varies with the seasons. Average NRMSE ($R^2$) values for February, August, and November are 0.30 (0.91), 0.32 (0.90), and 0.31 (0.90), respectively, while May has the highest average NRMSE of 0.34 and the lowest average $R^2$ of 0.88 among the four seasons, indicating poorer estimation accuracy in summer. The figure also shows higher NRMSE values at most in summer compared to other seasons.

Figure 11b illustrates the distribution of NRMSE and $R^2$ for the model-estimated SS across different seasons and depths. The NRMSE values show an increasing trend from the sea surface to 100 m, then level off between 200 m and 1200 m, and increase again at

depths greater than 1200 m. This reflects that the DO-ResNet model is able to effectively estimate subsurface salinity in the 1200 m depth layer based on surface data obtained from remote sensing, but its accuracy in estimating salinity at deeper depths is reduced. From the figure, it can be seen that the maximum NRMSE values for February and August occur around a depth of 70 m, being 0.42 and 0.35, respectively, while for May and November, the maximum values are 0.36 and 0.39, respectively. This indicates that the DO-ResNet model has larger errors in spring and autumn at 70 m, which may be affected by seasonal changes in freshwater fluxes [74] and the thermocline [75]. In general, the seasons had similar vertically averaged NRMSE and $R^2$, with NRMSE around 0.22 and $R^2$ around 0.94, which suggests that the DO-ResNet model is less seasonally influenced and robust in estimating SS.

### 4.5. Correlation Analysis between the ST (SS) and Surface Parameters

In order to evaluate how sea surface parameters influence the performance of the DO-ResNet model, we determined the Pearson correlation coefficients between SST, SSS, SSHA, USSW, and VSSW and the ST and SS at various depths (50 m, 60 m, 100 m, 200 m, 500 m, 1000 m, 1500 m, and 1900 m) presented in Figure 12. A higher absolute value of the Pearson correlation coefficient indicates a greater influence of the sea surface parameter on the estimation of ST or SS. By comparing these coefficients, we can determine the significance of each parameter in the model's estimation of ST and SS at different depths.
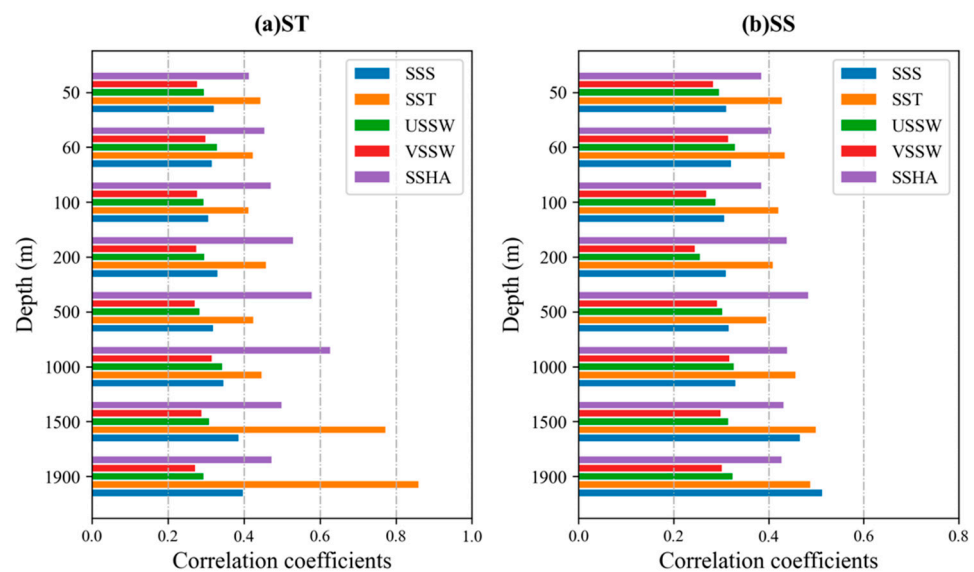


**Figure 12.** Pearson correlation coefficients between the DO-ResNet model estimated ST (**a**) and SS (**b**) at different depths and the observational sea surface parameters.

Figure 12a illustrates the Pearson correlation coefficients between sea surface parameters and the estimated ST. It is evident that SST and SSHA exhibit a strong correlation with ST. Specifically, SST has the highest correlation with ST at 50 m, with a coefficient of 0.86, while SSHA shows the highest correlation at 100 m, with a coefficient of 0.63. In contrast, USSW and VSSW display a weaker correlation with ST, with most coefficients being less than 0.33. SSS significantly impacts ST in the shallow layers, but its correlation diminishes with depth, reaching a minimum coefficient of 0.31 at 1000 m. Beyond 1000 m, its correlation is comparable to that of USSW and VSSW.

Figure 12b shows the correlation of various sea surface parameters with the DO-ResNet model's estimates of SS at different depths. The results indicate that SSS, SST, and SSHA have significant impacts on SS estimation. Specifically, SSS exhibits a high correlation with SS in the surface layer, with a correlation coefficient of 0.51 at 50 m. However, its correlation weakens in the deeper ocean, reaching approximately 0.31. SST and SSHA

maintain correlation coefficients above 0.4 at most depths, indicating their considerable and stable influence on SS estimation across different depth ranges. In contrast, USSW and VSSW exhibit weaker correlations with SS. Nevertheless, these parameters still contribute to SS estimation and should not be disregarded in the estimation process. The analysis indicates that the selected sea surface parameters are effective variables for estimating both ST and SS.

**5. Conclusions**

As the region with the highest sea surface temperatures globally, the tropical Western Pacific exhibits significant climate signals across various time scales. Its temperature and salinity demonstrate variations on interannual and decadal scales associated with major climate patterns such as the ENSO and the Pacific Decadal Oscillation. Accurately estimating the vertical structure of three-dimensional temperature and salinity fields in the tropical Western Pacific is crucial for understanding oceanic processes and climate change. However, due to the challenges and high costs associated with in situ observations, temperature and salinity data in the tropical Western Pacific remain sparse. This study introduces an improved DO-ResNet model, utilizing satellite-derived sea surface data (SSS, SST, SSHA, and SSW) and geographic information (longitude and latitude) as input, with Argo data employed as the ground truth for training and validation. The model reconstructs ST and SS in the tropical Western Pacific. The accuracy and reliability of the DO-ResNet model's estimates are assessed using RMSE, NRMSE, and $R^2$ metrics. Furthermore, the model's performance is evaluated from various perspectives, including spatial distribution, vertical sections, and seasonal variations.

This study evaluates the accuracy of the DO-ResNet model in estimating ST and SS by comparing its output at various depths with observational data from Argo floats. Results indicate that the optimized composite loss function enables the DO-ResNet model to effectively reconstruct most observed features of ST and SS concurrently in the tropical Western Pacific. However, due to the presence of the thermocline, the estimation accuracy of ST and SS is lower at depths around 150–200 m, posing challenges for accurately estimating the three-dimensional temperature and salinity fields near this depth. Despite this, the overall performance of the DO-ResNet model is commendable, with the average RMSE and $R^2$ for ST (SS) being 0.34 °C (0.05 psu) and 0.91 (0.95), respectively. Furthermore, this study evaluates the impact of sea surface parameters on the performance of the DO-ResNet model, indicating that sea surface variables play a more significant role in the upper ocean, particularly in shallow layers. SST and SSHA are crucial for estimating ST, while SSS, SST, and SSHA are key for estimating SS. Moreover, the model exhibits low NRMSE and high $R^2$ values across all seasons, indicating its robust seasonal applicability for ST and SS estimation in the tropical Western Pacific.

In summary, the DO-ResNet model demonstrates superior performance in estimating ST and SS in the tropical Western Pacific. The model exhibits high estimation accuracy for ST and SS across different depths and seasons. The results of this study are significant for understanding the subsurface information of the tropical Western Pacific in the context of ENSO occurrence and provide a foundation for further research into the ENSO mechanisms in the tropical Western Pacific. Future research should focus on employing more advanced artificial intelligence methods and integrating ocean dynamic mechanisms to further enhance estimation accuracy. Additionally, the application of this model can be extended to broader oceanic regions, allowing for precise estimation of ST and SS and facilitating practical applications such as mixed layer depth estimation and ocean hazard prediction. Moreover, the model can be further utilized to estimate other critical ocean parameters, such as velocity fields and ocean density, thus opening up new avenues for future research. It is also recommended to use TAO/TRITON data for cross-validation in future studies to investigate whether it can improve model performance and increase its credibility in practical applications.

## References

1. Smale, D.A.; Wernberg, T.; Oliver, E.C.J.; Thomsen, M.; Harvey, B.P.; Straub, S.C.; Burrows, M.T.; Alexander, L.V.; Benthuysen, J.A.; Donat, M.G.; et al. Marine Heatwaves Threaten Global Biodiversity and the Provision of Ecosystem Services. *Nat. Clim. Chang.* **2019**, *9*, 306–312. [CrossRef]
2. Du, Y.; Zhang, Y.; Zhang, L.-Y.; Tozuka, T.; Ng, B.; Cai, W. Thermocline Warming Induced Extreme Indian Ocean Dipole in 2019. *Geophys. Res. Lett.* **2020**, *47*, e2020GL090079. [CrossRef]
3. Planton, Y.Y.; Vialard, J.; Guilyardi, E.; Lengaigne, M.; McPhaden, M.J. The Asymmetric Influence of Ocean Heat Content on ENSO Predictability in the CNRM-CM5 Coupled General Circulation Model. *J. Clim.* **2021**, *34*, 5775–5793. [CrossRef]
4. Durack, P.J.; Wijffels, S.E.; Matear, R.J. Ocean Salinities Reveal Strong Global Water Cycle Intensification During 1950 to 2000. *Science* **2012**, *336*, 455–458. [CrossRef] [PubMed]
5. Stark, S.; Wood, R.A.; Banks, H.T. Reevaluating the Causes of Observed Changes in Indian Ocean Water Masses. *J. Clim.* **2006**, *19*, 4075–4086. [CrossRef]
6. Barreiro, M.; Fedorov, A.; Pacanowski, R.; Philander, S. Abrupt Climate Changes: How Freshening of the Northern Atlantic Affects the Thermohaline and Wind-Driven Oceanic Circulations. *Annu. Rev. Earth Planet. Sci.* **2008**, *36*, 33–58. [CrossRef]
7. Ghil, M.; Malanotte-Rizzoli, P. Data Assimilation in Meteorology and Oceanography. *Adv. Geophys.* **1991**, *33*, 141–266. [CrossRef]
8. Troccoli, A.; Haines, K. Use of the Temperature–Salinity Relation in a Data Assimilation Context. *J. Atmos. Ocean. Technol.* **1999**, *16*, 2011–2025. [CrossRef]
9. Vossepoel, F.C.; Behringer, D.W. Impact of Sea Level Assimilation on Salinity Variability in the Western Equatorial Pacific. *J. Phys. Oceanogr.* **2000**, *30*, 1706–1721. [CrossRef]
10. Carrassi, A.; Bocquet, M.; Bertino, L.; Evensen, G. Data Assimilation in the Geosciences: An Overview of Methods, Issues, and Perspectives. *Wiley Interdiscip. Rev. Clim. Chang.* **2018**, *9*, e535. [CrossRef]
11. Moore, A.M.; Martin, M.J.; Akella, S.; Arango, H.G.; Balmaseda, M.; Bertino, L.; Ciavatta, S.; Cornuelle, B.; Cummings, J.; Frolov, S.; et al. Synthesis of Ocean Observations Using Data Assimilation for Operational, Real-Time and Reanalysis Systems: A More Complete Picture of the State of the Ocean. *Front. Mar. Sci.* **2019**, *6*, 90. [CrossRef]
12. Kako, S.; Kubota, M. Numerical Study on the Variability of Mixed Layer Temperature in the North Pacific. *J. Phys. Oceanogr.* **2009**, *39*, 737–752. [CrossRef]
13. Wan, L.; Bertino, L.; Zhu, J. Assimilating Altimetry Data into a HYCOM Model of the Pacific: Ensemble Optimal Interpolation versus Ensemble Kalman Filter. *J. Atmos. Ocean. Technol.* **2010**, *27*, 753–765. [CrossRef]
14. Meng, L.; Yan, X.-H. Remote Sensing for Subsurface and Deeper Oceans: An Overview and a Future Outlook. *IEEE Trans. Geosci. Remote Sens.* **2022**, *10*, 72–92. [CrossRef]
15. Fiedler, P.C. Surface Manifestations of Subsurface Thermal Structure in the California Current. *J. Geophys. Res. Oceans* **1988**, *93*, 4975–4983. [CrossRef]
16. Chu, P.C.; Fan, C.; Liu, W.T. Determination of Vertical Thermal Structure from Sea Surface Temperature. *J. Atmos. Ocean. Technol.* **2000**, *17*, 971–979. [CrossRef]
17. Willis, J.K.; Roemmich, D.; Cornuelle, B. Combining Altimetric Height with Broadscale Profile Data to Estimate Steric Height, Heat Storage, Subsurface Temperature, and Sea-Surface Temperature Variability. *J. Geophys. Res. Ocean.* **2003**, *108*, 3292. [CrossRef]

18. Carnes, M.R.; Mitchell, J.L.; de Witt, P.W. Synthetic Temperature Profiles Derived from Geosat Altimetry: Comparison with Air-Dropped Expendable Bathythermograph Profiles. *J. Geophys. Res. Ocean.* **1990**, *95*, 17979–17992. [CrossRef]
19. Carnes, M.R.; Teague, W.J.; Mitchell, J.L. Inference of Subsurface Thermohaline Structure from Fields Measurable by Satellite. *J. Atmos. Ocean. Technol.* **1994**, *11*, 551–566. [CrossRef]
20. Yan, H.; Wang, H.; Zhang, R.; Chen, J.; Bao, S.; Wang, G. A Dynamical-Statistical Approach to Retrieve the Ocean Interior Structure From Surface Data: SQG-mEOF-R. *J. Geophys. Res. Ocean.* **2020**, *125*, e2019JC015840. [CrossRef]
21. Maes, C.; Behringer, D.; Reynolds, R.W.; Ji, M. Retrospective Analysis of the Salinity Variability in the Western Tropical Pacific Ocean Using an Indirect Minimization Approach. *J. Atmos. Ocean. Technol.* **2000**, *17*, 512–524. [CrossRef]
22. Guinehut, S.; Dhomps, A.-L.; Larnicol, G.; Le Traon, P.-Y. High Resolution 3-D Temperature and Salinity Fields Derived from in Situ and Satellite Observations. *Ocean Sci.* **2012**, *8*, 845–857. [CrossRef]
23. Baldacci, A.; Corsini, G.; Grasso, R.; Manzella, G.; Allen, J.T.; Cipollini, P.; Guymer, T.H.; Snaith, H.M. A Study of the Alboran Sea Mesoscale System by Means of Empirical Orthogonal Function Decomposition of Satellite Data. *J. Mar. Syst.* **2001**, *29*, 293–311. [CrossRef]
24. Zheng, G.; Li, X.; Zhang, R.-H.; Liu, B. Purely Satellite Data–Driven Deep Learning Forecast of Complicated Tropical Instability Waves. *Sci. Adv.* **2020**, *6*, eaba1482. [CrossRef] [PubMed]
25. Zhang, X.; Li, X. Combination of Satellite Observations and Machine Learning Method for Internal Wave Forecast in the Sulu and Celebes Seas. *IEEE Trans. Geosci. Remote Sens.* **2021**, *59*, 2822–2832. [CrossRef]
26. Zhang, X.; Li, X.; Zheng, Q. A Machine-Learning Model for Forecasting Internal Wave Propagation in the Andaman Sea. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2021**, *14*, 3095–3106. [CrossRef]
27. Wang, Y.; Li, X.; Song, J.; Li, X.; Zhong, G.; Zhang, B. Carbon Sinks and Variations of pCO2 in the Southern Ocean From 1998 to 2018 Based on a Deep Learning Approach. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2021**, *14*, 3495–3503. [CrossRef]
28. Ali, M.M.; Swain, D.; Weller, R.A. Estimation of Ocean Subsurface Thermal Structure from Surface Parameters: A Neural Network Approach. *Geophys. Res. Lett.* **2004**, *31*, L20308. [CrossRef]
29. Wu, X.; Yan, X.-H.; Jo, Y.-H.; Liu, W.T. Estimation of Subsurface Temperature Anomaly in the North Atlantic Using a Self-Organizing Map Neural Network. *J. Atmos. Ocean. Technol.* **2012**, *29*, 1675–1688. [CrossRef]
30. Chen, C.; Yang, K.; Ma, Y.; Wang, Y. Reconstructing the Subsurface Temperature Field by Using Sea Surface Data Through Self-Organizing Map Method. *IEEE Geosci. Remote. Sens. Lett.* **2018**, *15*, 1812–1816. [CrossRef]
31. Dong, L.; Qi, J.; Yin, B.; Zhi, H.; Li, D.; Yang, S.; Wang, W.; Cai, H.; Xie, B. Reconstruction of Subsurface Salinity Structure in the South China Sea Using Satellite Observations: A LightGBM-Based Deep Forest Method. *Remote Sens.* **2022**, *14*, 3494. [CrossRef]
32. Li, W.; Su, H.; Wang, X.; Yan, X. Estimation of Global Subsurface Temperature Anomaly Based on Multisource Satellite Observations. *J. Remote Sens.* **2017**, *21*, 881–891. [CrossRef]
33. Su, H.; Li, W.; Yan, X.-H. Retrieving Temperature Anomaly in the Global Subsurface and Deeper Ocean From Satellite Observations. *J. Geophys. Res. Ocean.* **2018**, *123*, 399–410. [CrossRef]
34. Su, H.; Wu, X.; Yan, X.-H.; Kidwell, A. Estimation of Subsurface Temperature Anomaly in the Indian Ocean during Recent Global Surface Warming Hiatus from Satellite Measurements: A Support Vector Machine Approach. *Remote Sens. Environ.* **2015**, *160*, 63–71. [CrossRef]
35. Su, H.; Yang, X.; Lu, W.; Yan, X.-H. Estimating Subsurface Thermohaline Structure of the Global Ocean Using Surface Remote Sensing Observations. *Remote Sens.* **2019**, *11*, 1598. [CrossRef]
36. Meng, L.; Yan, C.; Zhuang, W.; Zhang, W.; Geng, X.; Yan, X.-H. Reconstructing High-Resolution Ocean Subsurface and Interior Temperature and Salinity Anomalies From Satellite Observations. *IEEE Trans. Geosci. Remote Sens.* **2022**, *60*, 4104114. [CrossRef]
37. Su, H.; Zhang, T.; Lin, M.; Lu, W.; Yan, X.-H. Predicting Subsurface Thermohaline Structure from Remote Sensing Data Based on Long Short-Term Memory Neural Networks. *Remote Sens. Environ.* **2021**, *260*, 112465. [CrossRef]
38. Cheng, H.; Sun, L.; Li, J. Neural Network Approach to Retrieving Ocean Subsurface Temperatures from Surface Parameters Observed by Satellites. *Water* **2021**, *13*, 388. [CrossRef]
39. Mao, K.; Liu, C.; Zhang, S.; Gao, F. Reconstructing Ocean Subsurface Temperature and Salinity from Sea Surface Information Based on Dual Path Convolutional Neural Networks. *J. Mar. Sci. Eng.* **2023**, *11*, 1030. [CrossRef]
40. Wang, H.; Song, T.; Zhu, S.; Yang, S.; Feng, L. Subsurface Temperature Estimation from Sea Surface Data Using Neural Network Models in the Western Pacific Ocean. *Mathematics* **2021**, *9*, 852. [CrossRef]
41. Bao, S.; Zhang, R.; Wang, H.; Yan, H.; Yu, Y.; Chen, J. Salinity Profile Estimation in the Pacific Ocean from Satellite Surface Salinity Observations. *J. Atmos. Ocean. Technol.* **2019**, *36*, 53–68. [CrossRef]
42. Chang, L.; Xu, J.; Tie, X.; Wu, J. Impact of the 2015 El Nino Event on Winter Air Quality in China. *Sci. Rep.* **2016**, *6*, 34275. [CrossRef] [PubMed]
43. Zhai, P.; Yu, R.; Guo, Y.; Li, Q.; Ren, X.; Wang, Y.; Xu, W.; Liu, Y.; Ding, Y. The Strong El Niño of 2015/16 and Its Dominant Impacts on Global and China's Climate. *J. Meteorol. Res.* **2016**, *30*, 283–297. [CrossRef]
44. Cane, M.A. A Role for the Tropical Pacific. *Science* **1998**, *282*, 59–61. [CrossRef]
45. Picaut, J.; Ioualalen, M.; Menkes, C.; Delcroix, T.; McPhaden, M.J. Mechanism of the Zonal Displacements of the Pacific Warm Pool: Implications for ENSO. *Science* **1996**, *274*, 1486–1489. [CrossRef]
46. Matsuura, T.; Yumoto, M.; Iizuka, S. A Mechanism of Interdecadal Variability of Tropical Cyclone Activity over the Western North Pacific. *Clim. Dyn.* **2003**, *21*, 105–117. [CrossRef]

47.  Zhang, R.-H.; Zhou, G.; Zhi, H.; Gao, C.; Wang, H.; Feng, L. Salinity Interdecadal Variability in the Western Equatorial Pacific and Its Effects during 1950–2018. *Clim. Dyn.* **2023**, *60*, 1963–1985. [CrossRef]
48.  Huang, R.; Sun, F. Impacts of the Tropical Western Pacific on the East Asian Summer Monsoon. *J. Meteorol. Soc. Jpn. Ser. II* **1992**, *70*, 243–256. [CrossRef]
49.  Hackert, E.; Akella, S.; Ren, L.; Nakada, K.; Carton, J.A.; Molod, A. Impact of the TAO/TRITON Array on Reanalyses and Predictions of the 2015 El Niño. *J. Geophys. Res. Ocean.* **2023**, *128*, e2023JC020039. [CrossRef]
50.  Boutin, J.; Vergely, J.L.; Marchand, S.; D'Amico, F.; Hasson, A.; Kolodziejczyk, N.; Reul, N.; Reverdin, G.; Vialard, J. New SMOS Sea Surface Salinity with Reduced Systematic Errors and Improved Variability. *Remote Sens. Environ.* **2018**, *214*, 115–134. [CrossRef]
51.  Banzon, V.; Smith, T.M.; Chin, T.M.; Liu, C.; Hankins, W. A Long-Term Record of Blended Satellite and in Situ Sea-Surface Temperature for Climate Monitoring, Modeling and Environmental Studies. *Earth Syst. Sci. Data* **2016**, *8*, 165–176. [CrossRef]
52.  Hauser, D.; Tourain, C.; Hermozo, L.; Alraddawi, D.; Aouf, L.; Chapron, B.; Dalphinet, A.; Delaye, L.; Dalila, M.; Dormy, E.; et al. New Observations From the SWIM Radar On-Board CFOSAT: Instrument Validation and Ocean Wave Measurement Assessment. *IEEE Trans. Geosci. Remote Sens.* **2021**, *59*, 5–26. [CrossRef]
53.  Atlas, R.; Hoffman, R.N.; Ardizzone, J.; Leidner, S.M.; Jusem, J.C.; Smith, D.K.; Gombos, D. A Cross-Calibrated, Multiplatform Ocean Surface Wind Velocity Product for Meteorological and Oceanographic Applications. *Bull. Am. Meteorol. Soc.* **2011**, *92*, 157–174. [CrossRef]
54.  Roemmich, D.; Gilson, J. The 2004–2008 Mean and Annual Cycle of Temperature, Salinity, and Steric Height in the Global Ocean from the Argo Program. *Prog. Oceanogr.* **2009**, *82*, 81–100. [CrossRef]
55.  Wong, A.P.S.; Wijffels, S.E.; Riser, S.C.; Pouliquen, S.; Hosoda, S.; Roemmich, D.; Gilson, J.; Johnson, G.C.; Martini, K.; Murphy, D.J.; et al. Argo Data 1999–2019: Two Million Temperature-Salinity Profiles and Subsurface Velocity Observations From a Global Array of Profiling Floats. *Front. Mar. Sci.* **2020**, *7*, 700. [CrossRef]
56.  Rivera Tello, G.A.; Takahashi, K.; Karamperidou, C. Explained Predictions of Strong Eastern Pacific El Niño Events Using Deep Learning. *Sci. Rep.* **2023**, *13*, 21150. [CrossRef] [PubMed]
57.  He, K.; Zhang, X.; Ren, S.; Sun, J. Deep Residual Learning for Image Recognition. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Las Vegas, NV, USA, 27–30 June 2016; IEEE: Piscataway, NJ, USA, 2016.
58.  Chen, Z.; Jiang, Y.; Zhang, X.; Zheng, R.; Qiu, R.; Sun, Y.; Zhao, C.; Shang, H. ResNet18DNN: Prediction Approach of Drug-Induced Liver Injury by Deep Neural Network with ResNet18. *Brief Bioinform.* **2022**, *23*, bbab503. [CrossRef]
59.  Dzakmic, Š. Evaluation of ResNet Network for Semantic Segmentation of Coral Reefs. *Int. J. Eng. Technol.* **2020**, *8*, 54–61. [CrossRef]
60.  Hu, Y.; Hua, X.; Liu, W.; Wickert, J. Sea Ice Detection from GNSS-R Data Based on Residual Network. *Remote Sens.* **2023**, *15*, 4477. [CrossRef]
61.  Eshaq, R.M.A.; Hu, E.; Qaid, H.A.A.M.; Zhang, Y.; Liu, T. Using Deep Convolutional Neural Networks and Infrared Thermography to Identify Coal Quality and Gangue. *IEEE Access* **2021**, *9*, 147315–147327. [CrossRef]
62.  Yang, G.; Ye, X.; Xu, Q.; Yin, X.; Xu, S. Sea Surface Chlorophyll-a Concentration Retrieval from HY-1C Satellite Data Based on Residual Network. *Remote Sens.* **2023**, *15*, 3696. [CrossRef]
63.  Nair, V.; Hinton, G.E. Rectified Linear Units Improve Restricted Boltzmann Machines. In Proceedings of the 27th International Conference on International Conference on Machine Learning, Haifa, Israel, 21 June 2010.
64.  Terven, J.; Cordova-Esparza, D.M.; Ramirez-Pedraza, A.; Chavez-Urbiola, E.A. Loss Functions and Metrics in Deep Learning. *arXiv* **2023**, arXiv:2307.02694. [CrossRef]
65.  Wang, Q.; Ma, Y.; Zhao, K.; Tian, Y. A Comprehensive Survey of Loss Functions in Machine Learning. *Ann. Data. Sci.* **2022**, *9*, 187–212. [CrossRef]
66.  Talley, L.D.; Pickard, G.L.; Emery, W.J. *Descriptive Physical Oceanography: An Introduction*, 6th ed.; Academic Press: Amsterdam, The Netherlands; Boston, MA, USA, 2011; ISBN 978-0-7506-4552-2.
67.  Gueye, M.B.; Niang, A.; Arnault, S.; Thiria, S.; Crépon, M. Neural Approach to Inverting Complex System: Application to Ocean Salinity Profile Estimation from Surface Parameters. *Comput Geosci* **2014**, *72*, 201–209. [CrossRef]
68.  Su, H.; Lu, X.; Chen, Z.; Zhang, H.; Lu, W.; Wu, W. Estimating Coastal Chlorophyll-A Concentration from Time-Series OLCI Data Based on Machine Learning. *Remote Sens.* **2021**, *13*, 576. [CrossRef]
69.  Zhu, Y.; Zhang, R.-H.; Li, D.; Chen, D. The Thermocline Biases in the Tropical North Pacific and Their Attributions. *J. Clim.* **2021**, *34*, 1635–1648. [CrossRef]
70.  Qi, J.; Zhang, L.; Yin, B.; Li, D.; Xie, B.; Sun, G. Advancing Ocean Subsurface Thermal Structure Estimation in the Pacific Ocean: A Multi-Model Ensemble Machine Learning Approach. *Dyn. Atmos. Ocean.* **2023**, *104*, 101403. [CrossRef]
71.  Li, Y.; Wang, F. Thermohaline Intrusions in the Thermocline of the Western Tropical Pacific Ocean. *Acta Oceanol. Sin.* **2013**, *32*, 47–56. [CrossRef]
72.  Fine, R.A.; Lukas, R.; Bingham, F.M.; Warner, M.J.; Gammon, R.H. The Western Equatorial Pacific: A Water Mass Crossroads. *J. Geophys. Res.* **1994**, *99*, 25063–25080. [CrossRef]
73.  Qi, J.; Liu, C.; Chi, J.; Li, D.; Gao, L.; Yin, B. An Ensemble-Based Machine Learning Model for Estimation of Subsurface Thermal Structure in the South China Sea. *Remote Sens.* **2022**, *14*, 3207. [CrossRef]

74. Skliris, N.; Marsh, R.; Josey, S.A.; Good, S.A.; Liu, C.; Allan, R.P. Salinity Changes in the World Ocean since 1950 in Relation to Changing Surface Freshwater Fluxes. *Clim. Dyn.* **2014**, *43*, 709–736. [CrossRef]
75. Wang, L.; Xu, F. Decadal Variability and Trends of Oceanic Barrier Layers in Tropical Pacific. *Ocean Dyn.* **2018**, *68*, 1155–1168. [CrossRef]